Science Panel
August 25, 2010
Freshwater Sediment Standards
Review Questions

The Department of Ecology (Ecology) is engaged in a rule revision process of the Sediment Management Standards Chapter 173-204 WAC (SMS). Part of this rule revision process includes promulgation of both chemical and biological criteria for freshwater sediment. The SMS have chemical and biological criteria for marine sediment which were promulgated in 1991. These criteria are used to assess impacts to sediment benthic communities, identification of cleanup sites, and setting cleanup standards. Ecology is interested in promulgating criteria for freshwater sediment consistent with the current SMS scientific, technical, and policy framework.

Ecology is seeking independent peer-review on the science and methodology of both the chemical and biological criteria that have been proposed for promulgation. This peer review includes a technical report titled *"Development of Benthic SQGs for Freshwater Sediments in Washington, Oregon, and Idaho"* as well as the biological framework titled *"Overview of the Biological Freshwater Sediment Standards"*.

Reviewers will find it helpful to know that the proposed freshwater sediment criteria are consistent with the Washington State policy and scientific framework established in the SMS rule and promulgated in the SMS marine criteria. Specifically, the framework for the proposed freshwater sediment standards:

- Have two levels of chemical criteria. The lower level, *no adverse effects* (Sediment Quality Standard, SQS), serves as the goal for all sediment. The upper level, *minor adverse effects* (Cleanup Screening Level, CSL), serves as the screen above which a cleanup may be triggered.
- The lower criteria for each chemical is established as the lowest concentration with a biological hit from any one of the bioassays from the suite of bioassays.
- Have chemical criteria developed to maximize accuracy in predicting toxicity measured in bioassays (i.e., balances false positives and false negatives).
- The chemical criteria do not require cleanup based on proving absence of toxicity (for example, a single, conservative screening level set at a no adverse effects level or the lowest effects level).
- The cleanup standard for a particular cleanup site is established between the lower and upper levels (SQS and CSL), but as near as practical to the lower level based on net environmental benefits, technical feasibility, and cost.
- Have chemical criteria developed from regional data, screened at QA II level (litigation level) and ASTM and EPA or Ecology approved bioassays using organisms typical of the region.
- Have biological criteria that can be used to confirm impacts to the benthic community from chemical contaminants and can override chemistry results.
- Have biological criteria based on use of bioassays consistent with those used to develop the chemical criteria.

Reviewers will also find it helpful to know that Washington State regulations allow for biological test over-rides of chemical criteria which can be used when site specific conditions may result in reduced ability of the chemical criteria to predict toxicity. For example, if a sample resulted in no exceedances of the chemical criteria, but bioassays resulted in exceedances of the biological criteria, the sample would be considered an exceedance of the criteria (or a "hit"). Ecology will develop guidance that includes how to identify site specific conditions when the chemical criteria may not apply or may be under predictive of toxicity. These site specific conditions could include (but are not limited to): 1) water bodies not represented in the dataset used to derive the values, such as ephemeral/seasonal wetlands 2) metals driven mining sites 3) samples with unusual total organic carbon. For example high total organic carbon (TOC), TOC related to soot, coal tar, or, coal ash 4) unusual physical characteristics such as sites with wood waste, mining slag, etc. or 5) unusual water quality characteristics such as pH, or hardness.

REVIEW OF CHEMICAL CRITERIA

During the peer review of the chemical criteria, Ecology asks that reviewers focus on the following technical aspects of the technical report titled *"Development of SQGs for Freshwater Sediment in Oregon, Washington, and Idaho"*.

1. **General Approach:** Ecology has developed the draft sediment quality values by using a multivariate statistical model to characterize the relationship between chemical concentrations and sediment bioassay results from sediments collected from freshwater lakes, rivers, and streams in the Pacific Northwest.

    - Do you agree with Ecology's conclusion that the use of sediment bioassays provides a credible basis for predicting adverse benthic effects that is consistent with current scientific information?
    - Do you agree with Ecology's conclusion that multivariate statistical analysis provides a credible basis for characterizing the relationships between chemical concentrations and biological test results?

2. **Data Issues: The data used to develop the draft chemical criteria is summarized in Section 2.1 of the technical report.**

    - Do you agree with Ecology's conclusion that the database provides sufficient data to support the development of statewide chemical criteria? That is, is the dataset sufficiently robust for the development of chemical values specific to:
        - Geographical coverage.
        - Coverage of different types of freshwater systems.
        - Numbers of paired chemistry and bioassay endpoints.
        - Number of bioassay species.
        - Number of acute and chronic tests (referring to test duration relative to life history).
        - Number of lethal and sublethal effects endpoints.
        - Percentage of hits and no-hits for each endpoint.

- Do you agree with Ecology's conclusion that data was screened using criteria acceptable for the purposes of chemical criteria development? Specifically, refer to section 2.1.2 of the technical report which includes completeness of sediment chemical analysis, minimum number of detected analytes, QA/QC of sediment chemistry and bioassay data, and elimination of chemicals that were not directly associated with toxicity.

3. **Reliability Testing of the Chemical Criteria**: The state and federal agency representatives involved in this effort decided that it was preferred to use all available data to generate the criteria rather than to hold several datasets out of the process for validating the model. A validation study dependent upon new paired sediment chemistry and bioassay data would likely require years of data collection. Thus, reliability was evaluated with the dataset used to generate the values, looking at several reliability endpoints (refer to Section 2.3 Reliability Analysis in the technical report).

   - Do you believe that the approach used to evaluate the reliability of the criteria is consistent with current scientific methods and principles for validating criteria?
   - What comments do you have on the completeness and relative weight that should be given to the various reliability measures used to assess the results of the model and to compare it to other SQG sets?
   - Are there appropriate alternate validation methods that can use the data from which the standards were developed (i.e., bootstrapping methods, etc.)?
   - Comparative reliability analysis was used to assess different ways to handle data due to a number of issues. For example, petroleum compounds were reported using different analytical methods, (TPH vs. individual PAHs), and other similar issues occurred regarding use of the data (e.g., summing, normalization, inclusion of conventionals, splitting or lumping geographic areas). Given these differences, was the reliability comparison analysis performed for the development of these chemical criteria the most appropriate method or are there other methods to aid in selecting the best way to handle data?

4. **Data Interpretation and Use for Regulatory Decision-Making:**

   Greater differences in the chemical criteria occurred between the bioassays than between SQS and CSL level effects for any one bioassay endpoint. To select SQS and CSL levels for each chemical, the values for all bioassay endpoints and effects levels were combined into a single distribution representing the range of criteria from the lowest no adverse effects level to the highest minor adverse effects level. From this, the SQS was established as the lowest value and the CSL was selected as the next highest, significantly different value (see Table 3-7 of the technical report). This approach was selected as it provides conservative values by remaining at the low end of the no adverse effects to minor adverse effects distribution.

   - Do you agree with Ecology's conclusion that this approach is consistent with that of the WA SMS marine standards where the SQS and CSL were established as the lowest and 2nd lowest of the Apparent Effects Levels determined for a suite of bioassays?

- Given the types of historic sediment data available, is the TPH method the best available approach for assessing the overall effects of petroleum hydrocarbons? What is your answer based on (theory or empirical data)?
- Are the measures introduced in the model to assess covariance, coupled with other available statistical tests of covariance, sufficient to address the inevitable co-occurrence of chemicals in the field when developing chemical criteria?
- Should any chemical classes be summed that were not summed in the model to reduce covariance (e.g. phthalates)?

REVIEW OF BIOLOGICAL CRITERIA

While reviewing the biological criteria, Ecology asks that reviewers consider the technical and scientific aspects of using bioassays including bioassay organisms and endpoints. Please refer to the document titled *"Overview of biological freshwater sediment standards"*. The suite of bioassay species and endpoints were selected based partly on regional availability and familiarity with these organisms.

- Is the proposed bioassay suite appropriately sensitive to protect the freshwater macro benthic community (i.e., typical taxonomic structure and functions such as a prey base to endangered species like salmon)?
- From your experience, are there other freshwater bioassays/species that provide consistent, reproducible and sensitive results that should be considered for developing biological criteria?
- Are there issues you may be familiar with regarding running and interpreting these bioassays (e.g., problems associated with culturing animals, confounding variables that may warrant protocol modifications, choice of toxicant for positive controls, etc.)?
- What comments do you have on the appropriateness of the various growth endpoints under consideration nationally (e.g., dry weight, ash-free dry weight, length)?
- Is there additional information on the minimum detectable difference for these tests that would assist in setting the SQS (SL1) interpretive endpoint?
- What are the pros and cons of using two endpoints from the same bioassay (e.g., the Hyalella 28 day test has two endpoints commonly reported from the same test, mortality and growth). Are there ways to maximize the use of the combined results?